Recall: A set of vectors $S = \{\vec{x}_1, \dots, \vec{x}_k\} \subset \mathbb{R}^n$

is <u>orthogonal</u> if 1) $\vec{x}_i \neq \vec{0}_n$ for all i.

2) $\langle \vec{x}_i, \vec{x}_j \rangle = 0$ for $i \neq j$.

geometrically, this corresponds to being at right angles:

Recall: Theorem: Let $S = \{\vec{x}_1, \dots, \vec{x}_k\}$ be orthog.

Let $\vec{v} = \sum_{i=1}^{k} \alpha_i \vec{x}_i$, $\vec{u} = \sum_{j=1}^{k} \beta_j \vec{x}_j$.

Then $\langle \vec{v}, \vec{u} \rangle = \sum_{i=1}^{k} \alpha_i \beta_i \langle \vec{x}_i, \vec{x}_i \rangle$.

This theorem tells us that it is often useful to express vectors as linear combinations of orthogonal vectors.

Theorem: Suppose $\vec{v} \in \mathbb{R}^n$, $\{\vec{x}_1, \dots, \vec{x}_k\}$ orthogonal, & $\vec{v} = \sum_{i=1}^{k} \alpha_i \vec{x}_i$.

Then $\alpha_i = \dfrac{\langle \vec{x}_i, \vec{v} \rangle}{\langle \vec{x}_i, \vec{x}_i \rangle}$.

Remark: this theorem allows us to quickly compute α_i (much faster than using row reduction).

Proof: $\langle \vec{x}_i, \vec{v} \rangle = \langle \vec{x}_i, \sum_{j=1}^{k} \alpha_j \vec{x}_j \rangle$

k

$$= \sum_{j=1} \langle \vec{x}_i, \alpha_j \vec{x}_j \rangle \qquad \text{(distribute)}$$

$$= \sum_{j=1}^{k} \alpha_j \langle \vec{x}_i, \vec{x}_j \rangle$$

$$= \alpha_i \langle \vec{x}_i, \vec{x}_i \rangle \qquad \left(\text{since } \langle \vec{x}_i, \vec{x}_j \rangle = 0 \text{ for } i \neq j\right)$$

Hence $\dfrac{\langle \vec{x}_i, \vec{v} \rangle}{\langle \vec{x}_i, \vec{x}_i \rangle} = \alpha_i$ $\square$

Example: Let $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\quad \vec{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Claim: $\vec{x}_1, \vec{x}_2$ are orthogonal.

$$\langle \vec{x}_1, \vec{x}_2 \rangle = 1 \cdot 1 + 1(-1) = 0.$$

Let $\vec{v} = 3\vec{x}_1 + 2\vec{x}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

Let's check that our theorem gives the correct answers in this case:

$$\alpha_1 = \frac{\langle \vec{x}_1, \vec{v} \rangle}{\langle \vec{x}_1, \vec{x}_1 \rangle} = \frac{(1\cdot 5 + 1\cdot 1)}{(1\cdot 1 + 1\cdot 1)} = \frac{6}{2} = 3.$$

$$\alpha_2 = \frac{\langle \vec{x}_2, \vec{v} \rangle}{\langle \vec{x}_2, \vec{x}_2 \rangle} = \frac{(1\cdot 5 + (-1)\cdot 1)}{(1\cdot 1 + (-1)(-1))} = \frac{4}{2} = 2.$$

Orthogonal bases

Definition: An orthogonal basis of a subspace $V \subset \mathbb{R}^n$ is a basis of V which is also orthogonal.

Recall: a basis $S = \{\vec{x}_1, \dots, \vec{x}_k\}$ of V is a set of vectors such that:

1) S is linearly independent
2) $\langle S \rangle = V$.

Example: Let $V = \mathbb{R}^3$.

1) $\vec{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ $\vec{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $\vec{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

Exercise: check that this is an orthogonal basis.

2) $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ \end{bmatrix}$ $\vec{x}_2 = \begin{bmatrix} 1 \\ -1 \\ \end{bmatrix}$ $\vec{x}_3 = \begin{bmatrix} 0 \\ 0 \\ \end{bmatrix}$

Exercise: check that this is an orthogonal basis.

Theorem: Let $\vec{x}_1, \dots, \vec{x}_k$ be orthogonal. Then $\vec{x}_1, \dots, \vec{x}_k$ are linearly independent.

Proof: Suppose $\alpha_1 \vec{x}_1 + \dots + \alpha_k \vec{x}_k = \vec{0}_n$.

Want to show that $\alpha_1, \dots, \alpha_k = 0$.

$$\alpha_i = \frac{\left\langle \vec{x}_i, \sum_{j=1}^{k} \alpha_j \vec{x}_j \right\rangle}{\langle \vec{x}_i, \vec{x}_i \rangle}$$

by our previous theorem

$$= \frac{\langle \vec{x}_i, \vec{0}_n \rangle}{\langle \vec{x}_i, \vec{x}_i \rangle} = \frac{0}{\langle \vec{x}_i, \vec{x}_i \rangle} = 0. \checkmark$$

In our formulas, the expression $\langle \vec{x}_i, \vec{x}_i \rangle$ keeps popping up.
To simplify things, we make the following defi:

Definition: $\{\vec{x}_1, \dots, \vec{x}_k\} \subset \mathbb{R}^n$ is called <u>orthonormal</u>

if 1) $\{\vec{x}_1, \dots, \vec{x}_k\}$ is orthogonal

2) $\langle \vec{x}_i, \vec{x}_i \rangle = 1.$

orthogonal orthonormal

Remark: very similar to orthogonality, but simplifies lots of formulas.

Theorem: Let $V \subset \mathbb{R}^n$ be a subspace. Then V has an orthonormal basis.

To prove this theorem, one can write down an algorithm that produces such a basis:

<u>The Gram-Schmidt algorithm:</u>

Start with a basis of V, $b_1, \dots, b_k$.

We will modify $\vec{b}_1 \dots \vec{b}_k$ to produce an orthonormal basis of V, $\vec{x}_1, \dots, \vec{x}_k$.

1) $\vec{x}_1 = \dfrac{\vec{b}_1}{|\vec{b}_1|}$

2a) $\vec{x}_2' = \vec{b}_2 - \langle \vec{b}_2, \vec{x}_1 \rangle \vec{x}_1$

2b) $\vec{x}_2 = \dfrac{\vec{x}_2'}{|\vec{x}_2'|}$.

3a) $\vec{x}_3' = \vec{b}_3 - \langle \vec{b}_3, \vec{x}_1 \rangle \vec{x}_1 - \langle \vec{b}_3, \vec{x}_2 \rangle \vec{x}_2$

3b) $\vec{x}_3 = \dfrac{\vec{x}_3'}{|\vec{x}_3'|}$.

$\vdots$

j a) $\vec{x}_j{}' = \vec{b}_j - \sum_{i<j} \langle \vec{b}_j, \vec{x}_i \rangle \vec{x}_i$

j b) $\vec{x}_j = \dfrac{\vec{x}_j{}'}{|\vec{x}_j{}'|}$.

We thus obtain k new vectors $\vec{x}_1, \dots, \vec{x}_k$.

Theorem: $\vec{x}_1, \dots, \vec{x}_k$ is an orthonormal basis of V.

Proof: 1) We check $\langle \vec{x}_i, \vec{x}_j \rangle = 0$ for $i \neq j$.

Since $\langle \vec{x}_i, \vec{x}_j \rangle = \langle \vec{x}_j, \vec{x}_i \rangle$, we can assume $i < j$. We proceed by induction on j. Hence, suppose $\langle \vec{x}_i, \vec{x}_\ell \rangle = 0$ for $\ell < j$, $\ell \neq i$. Then:

$$\langle \vec{x}_i, \vec{x}_j \rangle = \left\langle \vec{x}_i, \frac{\vec{x}_j'}{|\vec{x}_j'|} \right\rangle = \frac{1}{|\vec{x}_j'|} \langle \vec{x}_i, \vec{x}_j' \rangle$$

$$= \frac{1}{|\vec{x}_j'|} \left\langle \vec{x}_i, \vec{b}_j - \sum_{\ell < j} \langle \vec{b}_j, \vec{x}_\ell \rangle \vec{x}_\ell \right\rangle$$

use inductive hypothesis:

$$= \frac{1}{|\vec{x}_j'|} \left(\langle \vec{x}_i, \vec{b}_j \rangle - \langle \vec{b}_j, \vec{x}_i \rangle \right) = 0.$$

Exercise: express $\vec{b}_j$ as a linear combination of the $\vec{x}_i$. Conclude that

$$\langle \vec{x}_1, \dots, \vec{x}_k \rangle = \langle \vec{b}_1, \dots, \vec{b}_k \rangle.$$

Exercise: Conclude that $\vec{x}_1, \dots, \vec{x}_k$ is an orthonormal basis of V.

Example: $V = \mathbb{R}^3$ $\quad \vec{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ $\quad \vec{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ $\quad \vec{b}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

Apply GS to $\vec{b}_1, \vec{b}_2, \vec{b}_3$:

$$\vec{x}_1 = \frac{\vec{b}_1}{|\vec{b}_1|} = \frac{\vec{b}_1}{\sqrt{1\cdot 1 + 1\cdot 1 \cdot 0 \cdot 0}} = \frac{\vec{b}_1}{\sqrt{2}} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}.$$

$$\vec{x}_2' = \vec{b}_2 - \langle \vec{b}_2, \vec{x}_1 \rangle \vec{x}_1$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \left(1\cdot\frac{1}{\sqrt{2}} + 1\cdot\frac{1}{\sqrt{2}} + 1\cdot 0\right)\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 1 \\ \end{bmatrix} - (2)\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \end{bmatrix}$$

$$\vec{x}_2 = \frac{\vec{x}_2'}{|\vec{x}_2'|} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

$$\vec{x}_3' = \vec{b}_3 - \langle \vec{b}_3, \vec{x}_1 \rangle \vec{x}_1 - \langle \vec{b}_3, \vec{x}_2 \rangle \vec{x}_2$$

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} - 0 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{bmatrix}$$

$$\vec{x}_3 = \frac{\vec{x}_3'}{|\vec{x}_3|} = \frac{\vec{x}_3'}{\sqrt{\left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2}} = \frac{\vec{x}_3'}{\sqrt{\frac{1}{2}}} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}.$$

Remark: To prove that the GS algorithm really does produce a basis, need only check that each $\vec{b}_i$ is a lin. combination of $\vec{x}_1, \dots, \vec{x}_k$.